DOCUMENT RESUME

ED 182 322 TM 010 002

AUTHOR Rodrigues, Arcldo
TITLE Analysis of Large Data Sets in Brazil: A Study on Selection Procedures to College Admission.
PUB DATE [Jul 79]
NOTE 13p.; Paper presented at the Annual Meeting of the Interamerican Congress of Psychology (17th, Lima, Peru, July 1-6, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Abstract Reasoning; Academic Ability; *Academic Aptitude; Achievement Tests; *Aptitude Tests; *College Entrance Examinations; College Majors; Foreign Countries; *Grade Point Average; *Grade Prediction; Higher Education; Nonverbal Ability; *Predictive Ability (Testing); Test Reliability; Test Validity; Verbal Ability
IDENTIFIERS Abstract Reasoning Test (Myers); Brazil

ABSTRACT

Because the achievement battery traditionally administered to Brazilian college bound students may discriminate against those who cannot afford to attend the best high schools or preparatory courses, tests were developed to measure aptitude for higher education, independent of ability or actual knowledge of course content. The focus of the research was the ability of two Brazilian-developed tests of verbal and numerical aptitude and the Myers Abstract Reasoning Test to predict grade point average after the first and fifth college semester. Results, based on 1,793 students, were reported by academic area: technology, biomedicine, and humanities. Although the tests were reported as being valid, reliable, and clear measures of some factor other than ability, their correlation with grades was not convincing. The tests were fair predictors of academic success in technology, moderate to poor in humanities, and very poor in biomedicine. The decreasing reliability of grade point averages was one of the reasons offered for these findings. Nevertheless, the trends found in the results, as well as the rationale behind aptitude tests, justify their refinement as college admission criteria. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY Aroldo Rodrigues TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED182322

TM010 002

ANALYSIS OF LARGE DATA SETS IN BRAZIL : A Study on Selection Procedures to College Admission.

Aroldo Rodrigues
Federal University of R. Grande do Sul
Faculty of Education

I - Introduction -

The present paper reports a work done by the CESGRANRIO Foundation, a Brazilian outfit somewhat similar to the College Board in the United States, regarding the usefulness of adding aptitude tests to the existing set of achievement tests used for selection of candidates to universities in the Great Rio de Janeiro area. It constitutes an example of the analysis of large data sets carried out in Brazil, in regard to educational matters.

The admission's system followed by the Foundation consists of submitting the candidates to schools of higher education to a series of achievement tests. These tests cover the course content of the disciplines taught during high-school. The candidate indicates his or her major areas as well as his or her particular schools of interest, in rank order up to three areas and schools. According to their raw scores converted into standard scores with mean 500 and standard deviation 100, the candidates are placed in the course and school that their average score, in competition with that of the other candidates, allows.

The problem felt by the Foundation was that this system may be unfair to those candidates who, owing to finantial limitations, cannot afford to attend the best high-schools and preparatory courses, thus having their chances of success severely impaired.

One way of reducing this handicap of those with finantial problems that forced them to go to high-schools of lower

quality and to, sometimes, work during the day and to study in the evenings, could be the utilization of aptitude tests in place of, or in addition to, the current used achievement tests.

This endeavour presents, of course, a variety of problems and it would take too much time to consider them here. Therefore, I singled out only one - the difficulty in distinguishing aptitude and ability - inasmuch as this problem is intimately related to the whole work to be reported here. I will not consider, therefore, problems such as whether or not only achievement tests covering high-school content are or not the best way of selecting candidates to universities; which other information about the candidates are needed for a wiser selection; etc..

When this research begun we were persuaded that the whole enterprise could only be meaningful if we could differentiate between aptitude and achievement tests. This distinction is not clear as we can see in the booklet in which the College Board describes the SAT to the candidate. There we see several references to the fact that the verbal and the numerical parts of the SAT will test the student's ability. Moreover, it is clearly said to the student that university officials need a standard measure of their ability since the various high-schools differ widely in their curricula, academic standards and grading systems.

We can clearly see in the above reference that the College Board calls Scholastic Aptitude Test a test in which ability seems to be the construct measured. We feel, however, that the two constructs can be differentiated. Aptitude is potential ability and ability is actualized potential. Two people, for example, can begin piano lessons at the same time, with the same

instructor, dedicating the same number of hours daily to practice, and, after 1 year, one plays clearly better than the other. We conclude that the better player has more aptitude than the other to play piano. On the other hand, two pianists may be roughly equivalent, but one studies much harder than the other; their abilities can be considered the same but their aptitudes are clearly different.

If this distinction is correct and operationalizeable, we should be able to come up with instruments that would detect potentiality (aptitude) for higher education independently of ability or actual knowledge of course content. We tried in this research to do just that.

We felt that there is general agreement as to the need for verbal, numerical and abstract reasoning in higher education. We then constructed two tests, one to measure verbal ability and one to measure numerical ability[1]. To measure abstract reasoning we were granted permission by the Educational Testing Service to use a test of abstract reasoning constructed by Charles T. Myers.

The verbal instrument followed the common format of such types of tests, care being taken to use only very common words found in the popular newspapers; the numerical test required only the knowledge of the four basic operations and the notion of percentage. Myers' abstract reasoning test was a non-verbal test, with

(1) Both tests were ~~concerned~~ constructed by the late João Telles Villas-Bôas Fº, who cooperated with the author in the selection of the tests.

20 pictorial stimuli only. We hoped, therefore, that a minimum amount of ability was necessary to succeed in the tests, and that aptitude for efficiently dealing with verbal, numerical and abstract materials was the key factor in the correct response to the items.

With the purpose and the instruments thus far briefly described, we set out to do the research which will now be described.

II - Method -

a) Subjects - 5 614 candidates to the entrance examination organized by the CESGRANRIO Foundation were submitted to the three aptitude tests mentioned above in 1973, 1974 and 1976. They constituted three different samples of those who, in these three years, took the entrance examination. Out of these 5 614, 1 949 were applying to the technological area, 1 667 to the biomedical sciences area, and 1 998 to humanities.

b) Procedure - The subjects were randomly selected (following a systematic random sample procedure) when they came to the Foundation to attain their authorization cards which would allow them to take the exams. 126 trained psychology students administered the tests to the total sample. The tests were scored and their raw scores converted into T scores with mean 500 and standard deviation 100. All subjects took, a month later, the examination which selected the more able to the several higher-education courses existing in Rio and affiliated to the Foundation.

The tests were administered in the following order: verbal (10'); numerical (30'); and abstract (30'). Between the

applications, there was an interval of 5 minutes for the subjects to rest.

In order to have a criterion of success in school, the grade point averages of those who succeeded in the exam and entered the university were obtained for each semester. For the sample of 1973 we had 7 semester grade point averages, for the 1974 sample we had 5; and for the 1976 we had only 1.

The sample of 5 614 subjects was reduced to 1 793 after the results of the entrance examination and after some drop-outs during the years that the research lasted.

III - Results -

Although the tests had been submitted to the usual verifications of their reliability and validity before, we report their coeficients of internal consistent found with the 1973 sample, in the three general areas - technological, biomedical and humanities, using formula 20 of Kuder-Richardson.

Table 1

Coefficients of Internal Consistency of the Three Tests

Areas	Tests		
	Verbal	Numerical	Abstract
Technological	.84	.78	.75
Biomedical	.85	.74	.68
Humanities	.81	.56	.65

The scores of both aptitude and achievement tests, with this same sample, were factor-analyzed and two factors clearly emerged. Table 2 shows the results of this analysis.

(Insert Table 2 here)

The validity of Meyers' Abstract Reasoning Test had already been established and that of the newly developed instruments - the verbal and the numerical - was determined using a concurrent validation procedure previous to their use in this research. Groups of university students in humanities and in the technological areas took the tests and the tests clearly and significantly differentiate between then in the expected direction. Reliability coefficients in the eighties were obtained for both using the split-half technique with Spearman-Brown correction.

The results to be reported next deal with the main point of this research, namely, whether or not the aptitude tests can predict school performance and thus contribute to the selection procedure.

Table 3 presents the correlations between scores on the tests and grade point average, corrected for restriction of range, in the three general areas for the samples of 73 and 74[1].

[1] The 1976 was small and, after the loss of subjects due to failure in the entrance examination and drop-outs, the number was reduced to meaningless figures.

Table 2

Factor Loadings and Communalities of 7 Variables
(4 achievement and 3 aptitude tests)

Technological Area				Biomedical Area				Humanities Area			
Variables	Factor 1: Ability	Factor 2: Aptitude	h^2	Variables	Factor 1: Ability	Factor 2: Aptitude	h^2	Variables	Factor 1: Ability	Factor 2: Aptitude	h^2
Languages	.524	.408	.441	Languages	.594	.344	.472	Portuguese	.678	.372	.598
Chemistry	.804	.248	.708	Biology	.810	.249	.718	Languages	.649	.237	.477
Physics	.812	.232	.714	Chemistry	.793	.278	.706	History	.772	.185	.631
Mathematics	.780	.152	.632	Physics + Mathematics	.698	.252	.552	Mathematics + Geography	.649	.372	.559
Verbal	.094	.583	.348	Verbal	.248	.535	.348	Verbal	.354	.421	.303
Numerical	.539	.514	.555	Numerical	.376	.516	.408	Numerical	.271	.600	.434
Abstract	.206	.544	.338	Abstract	.126	.559	.328	Abstract	.152	.607	.392

Table 3

Correlations between Predictor and Criterion Variables in the Three General Areas in Different Number of Semesters for Three Different Samples

	Technological				Biomedical				Humanities			
Predictors	'73		'74		'73		'74		'73		'74	
	1sem.	7sem.	1sem.	7sem.	1sem.	7sem.	1sem.	7sem.	1sem.	7sem.	1sem.	7sem.
Verbal	.23	.16	.13	.08	.25	-.02	-.14	-.11	.20	-.07	.15	.05
Numerical	.07	.14	.29	.26	-.15	.11	.15	.21	.04	.22	.11	.08
Abstract	.20	.34	.23	.23	.37	.22	-.09	-.07	.25	-.16	.07	.14
Σ Testes	.17	.32	.28	.22	.20	.15	-.05	.01	.21	.01	.10	.10
N	76	51	332	241	50	28	147	126	90	57	407	238

Further analyses were carried out with the entire sample, using the extreme groups approach. We obtained the distributions of scores in each aptitude test as well as in the sum of all three tests and also the distributions of grade point averages in the first semester and after five consecutive semesters (mean values). We dicotomized the sample using as cut-off points +1 standard deviation and -1 standard deviation, for all distributions of predictor variables (tests) and criterion variables (grade point averages). Table 4 shows the results found when means on criterion variables for high and low groups in the predictor variables were compared.

Table 4

Mean Criterion Values for High and Low Groups in the Predictor Variables for the Three General Areas

Predictor Variables	Groups	Technological			Biomedical			Humanities	
		N	1sem.	5sem.	N	1sem.	5sem.	N	1sem.	5sem.
Verbal	$\bar{X}$ + 1σ	41	6.41	6.19	39	6.35	6.59	44	7.08	6.89
	$\bar{X}$ - 1σ	35	5.65	5.83	29	6.32	7.02	78	6.60	6.78
	t		1.72	.78		.06	-.87		1.78	.34
	p		<.10	n.s.		n.s.	n.s.		n.s.	n.s.
Numerical	$\bar{X}$ + 1σ	47	6.41	6.47	43	6.97	7.07	43	7.13	7.16
	$\bar{X}$ - 1σ	35	4.59	5.55	26	6.37	7.04	65	6.25	6.85
	t		4.09	2.03		1.17	.07		2.64	.93
	p		<.001	<.05		n.s.	n.s.		<.01	n.s.
Abstract	$\bar{X}$ + 1σ	64	6.28	6.42	40	6.70	6.97	40	7.19	6.96
	$\bar{X}$ - 1σ	45	4.97	5.40	29	6.51	6.96	58	6.25	6.39
	t		3.63	2.95		.43	.03		2.74	.93
	p		.001	.004		n.s.	n.s.		<.01	n.s.
Σ Tests	$\bar{X}$ + 1σ	55	6.36	6.71	36	6.49	6.88	35	7.04	6.69
	$\bar{X}$ - 1σ	31	4.86	5.22	23	6.27	7.05	68	6.37	6.44
	t		3.12	3.65		.36	-.37		2.09	.65
	p		<.002	<.001		n.s.	n.s.		<.04	n.s.

Note: Ns for 5 semesters slightly different due to drop-outs.

In addition to this, tetrachoric r_s were calculated for predictor and criterion variables, dicotomized at + and - 1σ. Table 5 depicts the results.

Table 5

Tetrachoric Correlations between Dicotomized Predictor and Criterion Variables for the Three General Areas

Variables	r_t Technological	r_t Biomedical	r_t Humanities
Verbal x 1 semester	.58	.35	.46
Verbal x 5 semesters	.55	-.27	.39
Numerical x 1 semester	.51	.56	.10
Numerical x 5 semesters	.45	-.23	.16
Abstract x 1 semester	.76	.55	.56
Abstract x 5 semesters	.58	.09	.23
Σ Tests x 1 semester	.48	.33	.28
Σ Tests x 5 semesters	.76	.35	-.30

IV - Discussion -

The obtained results concerning the predictive value of the aptitude tests are far from convincing. Although the tests as such seem to be in good standing and despite the fact that they indeed tap something other than ability, the main objective of the research - i.e., the determination of the predictive value of these instruments - was not satisfactorily attained. The tests showed to be fair predictors of academic success in the technological area, moderate to bad predictors in the Humanities area, and very bad predictors in the Biomedical area. Let us see some possible reasons for these findings.

a) the unsuccessful results in the medical area can be due to the severe selection rate of this area. Only 20% of these candidates are admitted to the universities. The consequent huge restriction of range derived from this may have turned things difficult for the tests as far as discrimination is concerned;

b) grade point averages are becoming less and less reliable as a criterion measure; we have reason to believe that this criterion is least bad in the technological area (exactly the one where the tests showed some discriminative power);

c) a variety of institutional problems led to a series of compromises that were deleterious to the precision of the research endeavour. There seems to be no doubt that should the conditions have been more favorable, more consistent and clear-cut results would have emerged;

d) notwithstanding the fuziness of the results, it seems wise to continue to pursue the goal of refining aptitude tests to be used in the selection procedure of candidates to higher education along the general lines that inspired this work. The trends found in the data here reported, as well as the rationale behind this endeavour, seem to justify this assertion.